Topic 13
Method of Moments
Introduction and Procedure

Outline

Introduction

Method of moments estimation is based solely on the law of large numbers,

Let $M_1, M_2, \ldots$ be independent random variables having a common distribution possessing a mean μ_M. Then the sample means converge to the distributional mean as the number of observations increase.

$$\bar{M}_n = \frac{1}{n}\sum_{i=1}^{n} M_i \to \mu_M \quad \text{as } n \to \infty.$$

In addition, if the random variables in this sequence *fail* to have a mean, then the limit will *fail* to exist.

Procedure

- Step 1. If the model is based on a parametric family of densities $f_X(x|\theta)$ with a d-dimensional parameter space $(\theta_1, \theta_2 \ldots, \theta_d)$, we compute

$$\mu_m = EX^m = k_m(\theta) = \int_{-\infty}^{\infty} x^m f_X(x|\theta)\, dx, \quad m = 1, \ldots, d$$

  the first d moments,

$$\mu_1 = k_1(\theta_1, \theta_2 \ldots, \theta_d), \quad \mu_2 = k_2(\theta_1, \theta_2 \ldots, \theta_d), \quad \ldots, \quad \mu_d = k_d(\theta_1, \theta_2 \ldots, \theta_d),$$

  obtaining d equations in d unknowns.

Example

Let $X_1, X_2, \ldots, X_n$ be a simple random sample of Pareto random variables with density

$$f_X(x|\beta) = \frac{\beta}{x^{\beta+1}}, \quad x > 1, \qquad \beta > 0.$$

The cumulative distribution function is

$$F_X(x) = 1 - x^{-\beta}, \quad x > 1.$$

The mean and the variance are, respectively,

$$\mu = \frac{\beta}{\beta - 1}, \qquad \sigma^2 = \frac{\beta}{(\beta-1)^2(\beta-2)}.$$

In this situation, we have one parameter, namely β. Thus, in step 1, we will only need to determine the first moment

$$\mu_1 = \mu = k_1(\beta) = \frac{\beta}{\beta - 1}$$

to find the method of moments estimator $\hat{\beta}$ for β.

Procedure

- Step 2. We then solve for the d parameters as a function of the moments.

$$\theta_1 = g_1(\mu_1, \mu_2, \cdots, \mu_d), \quad \theta_2 = g_2(\mu_1, \mu_2, \cdots, \mu_d),$$
$$\ldots, \quad \theta_d = g_d(\mu_1, \mu_2, \cdots, \mu_d).$$

- Step 3. Now, based on the data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, we compute the first d sample moments,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2, \quad \ldots, \quad \overline{x^d} = \frac{1}{n}\sum_{i=1}^{n} x_i^d.$$

Example

Exercise. If

$$\mu = \frac{\beta}{\beta - 1}, \text{ show that } \beta = \frac{\mu}{\mu - 1}.$$

For step 2, we solve for β as a function of the mean μ.

$$\beta = g_1(\mu) = \frac{\mu}{\mu - 1}.$$

For step 3, we compute the sample mean

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Procedure

- Step 4. Using the law of large numbers, we have, for each moment, $m = 1, \ldots, d$, that

$$\mu_m \approx \overline{x^m}.$$

For the equations derived in step 2, we replace the distributional moments μ_m by the sample moments $\overline{x^m}$ to give us formulas for the method of moment estimates

$$(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_d).$$

For the data $\mathbf{x}$, these estimates are

$$\hat{\theta}_1(\mathbf{x}) = g_1(\bar{x}, \overline{x^2}, \cdots, \overline{x^d}), \quad \hat{\theta}_2(\mathbf{x}) = g_2(\bar{x}, \overline{x^2}, \cdots, \overline{x^d}),$$
$$\ldots, \quad \hat{\theta}_d(\mathbf{x}) = g_d(\bar{x}, \overline{x^2}, \cdots, \overline{x^d}).$$

Example

Consequently, a method of moments estimate for β is obtained by replacing the distributional mean μ by the sample mean $\bar{x}$ in the equation for $g_1(\mu)$. Thus,

$$\hat{\beta} = \frac{\bar{x}}{\bar{x} - 1}.$$

A good estimator should have a small variance. To use the delta method to estimate the variance of $\hat{\beta}$,

$$\sigma^2_{\hat{\beta}} \approx g_1'(\mu)^2 \frac{\sigma^2}{n}.$$

We compute

$$g_1(\mu) = \frac{\mu}{\mu - 1} \quad \text{and so} \quad g_1'(\mu) = -\frac{1}{(\mu - 1)^2}$$

giving

$$g_1'\left(\frac{\beta}{\beta - 1}\right) = -\frac{1}{(\frac{\beta}{\beta-1} - 1)^2} = -\frac{(\beta - 1)^2}{(\beta - (\beta - 1))^2} = -(\beta - 1)^2.$$

Example

We find that $\hat{\beta}$ has mean approximately equal to β and variance

$$\sigma^2_{\hat{\beta}} \approx g_1'(\mu)^2 \frac{\sigma^2}{n} = (\beta-1)^4 \frac{\beta}{n(\beta-1)^2(\beta-2)} = \frac{\beta(\beta-1)^2}{n(\beta-2)}.$$

Let's consider the case with $\beta = 4$ and $n = 225$. Then,

$$\sigma^2_{\hat{\beta}} \approx \frac{4 \cdot 3^2}{225 \cdot 2} = \frac{36}{450} = \frac{2}{25}, \quad \sigma_{\hat{\beta}} \approx \frac{\sqrt{2}}{5} = 0.283.$$

To simulate, we use the probability transform

$$u = F_X(x) = 1 - x^{-\beta}, \quad \text{then} \quad x = (1-u)^{-1/\beta} = 1/\sqrt[\beta]{(1-u)}.$$

Note that if U_i are $U(0,1)$ random variables, then $1/\sqrt[\beta]{(1-U_1)}, 1/\sqrt[\beta]{(1-U_2)}, \cdots$ have the appropriate Pareto distribution.

Example

```
> paretobar<-rep(0,1000)
> for (i in 1:1000){u<-runif(225);
   pareto<-1/(1-u)^(1/4);
   paretobar[i]<-mean(pareto)}
> betahat<-paretobar/(paretobar-1)
> mean(betahat)
[1] 4.03508
> sd(betahat)
[1] 0.2833142
```

Note that the mean is above 4, but the standard deviation is very close to the value given by the delta method.

Exercise. Reproduce the simulation above and compare. Simulate using a different value for β.

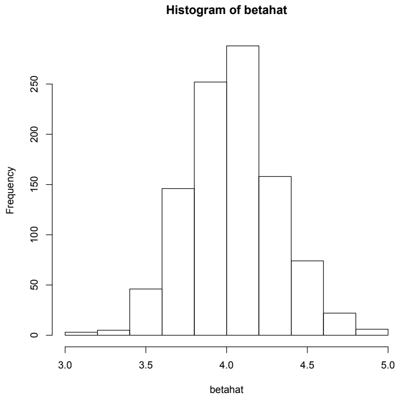

Figure: 1000 simulations for the method of moments estimate for the case $\beta = 4$.